DOCUMENT RESUME

ED 331 924 UD 028 025

TITLE	Summary Report on the Results of the April 1990 Standardized Testing Program. Research, Evaluation, and Planning Report.
SPONS AGENCY	Chicago Public Schools, Ill.
PUB DATE	Jan 91
NOTE	28p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Academic Standards; Educational Change; Elementary Secondary Education; *National Norms; Program Implementation; *Standardized Tests; *Testing Programs; Test Interpretation; Test Norms; *Test Results; Test Use
IDENTIFIERS	*Chicago Public Schools IL; Iowa Tests of Basic Skills; Tests of Achievement and Proficiency

ABSTRACT

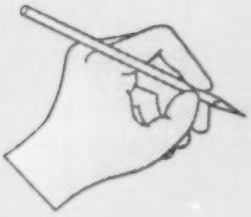
The Chicago (Illinois) School Reform Act of 1988 (Public Act 85-1418) established the performance of students on a nationally normed, standardized test as a key measure of success. In the Chicago Public Schools, the tests used are the Iowa Tests of Basic Skills (ITBS) in the elementary grades and the Tests of Achievement and Proficiency (TAP) in high school. This report summarizes the results achieved on the ITBS and TAP testing conducted in spring 1990, during the first year of the reform implementation. This information is seen as establishing a baseline against which progress can be measured during the next several years. This baseline can be expressed in several ways, but the simplest is the percentage of students below the national norms. The key tables in the report present the information in this format. This measure provides a convenient way of knowing how much progress is being made toward the legislative goal of having at least half of all students in each school at or above national norms, regardless of race, ethnicity, gender, or family income. Results indicate that, depending upon the grade and subject, improvement of 20 to 30 percent or more must take place. Eight tables of statistical data are included. (AF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED331924

REPORT

RESEARCH, EVALUATION & PLANNING



SUMMARY REPORT ON THE RESULTS OF THE APRIL 1990 STANDARDIZED TESTING PROGRAM

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

F. Schuster

Chicago Board of Ed

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

UD 028 025

BEST COPY AVAILABLE

CHICAGO PUBLIC SCHOOLS

Ted D. Kimbrough
General Superintendent of Schools

SUMMARY REPORT ON THE RESULTS OF THE APRIL 1990 STANDARDIZED TESTING PROGRAM

Prepared by

Department of Research, Evaluation and Planning

Ted D. Kimbrough
General Superintendent of Schools

January 1991

BOARD OF EDUCATION OF THE CITY OF CHICAGO

Clinton Bristow, Jr., President
Pamela A. Lenane, Vice President

Grady Bailey, Jr.
Stephen R. Ballis
Saundra Bishop
Florence B. Cox
Juan S. Cruz
Patricia A. Daley
The Rev. Darryl F. James
Rev. Nathaniel Jarrett
Albert N. Logan
Bertha G. Magaña
Anna Mustafa
Ashish Sen
Maria J. Vargas

It is the policy of the Board of Education of the City of Chicago not to discriminate on the basis of race, color, creed, national origin, religion, age, handicap unrelated to ability, or sex in its educational program or employment policies or practices.

Inquiries concerning the application of Title IX of the Education Amendments of 1972 and the regulations promulgated thereunder concerning sex discrimination should be referred to the Title IX Coordinator, Board of Education of the City of Chicago, 1819 West Pershing Road, Chicago, Illinois 60609.

CONTENTS

Introduction	1
The Testing Program	4
Summary of the Results on the *ITBS*	6
Summary of the Results on the *TAP*	13
Interpreting School Level Test Results	17
Future Reporting of Test Results	21

Educators can and do use various measures for determining the progress of students and the effect of educational programs. These include standardized tests, criterion-referenced tests, teacher-designed tests, performance assessments, and portfolios of students' work. All are useful; none completely describe the educational progress that has occured.

The Chicago School Reform Act of 1988 (P.L. 85-1418) establishes as one of its key measures of success the performance of students on a nationally normed, standardized test. In the Chicago Public Schools, the tests used are the *Iowa Tests of Basic Skills* (ITBS) in the elementary grades and the *Tests of Achievement and Proficiency* (TAP) in high school.

This report summarizes the results achieved on the ITBS and TAP testing conducted last spring, during the first year of the reform implementation. This information establishes a baseline against which progress can be measured during the next several years. This baseline can be expressed in several ways -- percentile ranks, grade-equivalent scores, etc. -- but the simplest is the percentage of students below the national norms. The key tables in this report present the information in this format.

This measure -- albeit somewhat simplified -- provides a convenient way of knowing how much progress is being made toward the legislative goal of having at least half of all students in each school at or above national norms, regardless of race, ethnicity, gender or family income. As can be seen from these tables, significant improvement must occur: depending upon the grade and subject (and the racial/gender classification), changes of 20 to 30 percent or more must take place.

For the first time, the Chicago Public Schools' results are being reported -- both to the schools and in this systemwide summary -- with breakdowns by students' gender and by racial/ethnic background. This information is presented to assist schools in diagnosing problems and needs and in suggesting solutions. These breakdowns, however, are as subject to misinterpretation as are other test results. Presenting scores for subgroups is not intended to imply, nor does it imply, that the results are caused by the characteristic of the subgroup. The data should not be read, for example, to conclude that girls score below boys in math because they are

girls. Where such differences between groups are found, they suggest that some other conditions are at work. The challenge is to identify those conditions and change them so that all subgroups perform excellently.

This is where the Chicago school system is at the beginning of reform. It is the chalk line drawn to mark the starting point. In subsequent years, further reports will show how far the system has progressed beyond that line.

THE TESTING PROGRAM

An important part of education is testing. It tells us whether or not students have learned what they were expected to learn. It also helps us to determine how good our educational programs and our instruction strategies are. It permits teachers and principals to find the areas where particular students excel and where they need more help. Finally, it serves as a key measure of the success of the school and of the school system in accomplishing their missions.

There are many ways to test students. In recent years, more and more attention and emphasis has been given to what is called 'norm-referenced standardized testing'. This is, in simple terms, a series of tests that is used by many school districts around the country, and that compares students' performance to the performance of a representative national sample of students.

Norm-referenced tests are extremely useful for telling us how well a student, or a group of students, or a school is doing *as compared to* other students or schools. Results on norm-referenced tests can be reported in various ways: as percentiles, as grade-equivalents, as stanines. But regardless of how they are reported, the scores are comparative measures. We know from the test results how our students did, compared to other students. We do not know from norm-referenced tests how much our students learned, compared to what they were taught.

The Chicago Public Schools uses two -- related -- sets of standardized, norm-referenced tests: the *Iowa Tests of Basic Skills* (ITBS) in the elementary schools, and the *Tests of Achievement and Proficiency* (TAP) in the high schools.

Although versions of these tests have been used for some years in Chicago, the particular tests given in April 1990 were new. This was done for two reasons: first, the old tests were based on norms established in 1978 which reduced the value of comparisons being made; second, the old tests had been in use for many years and teachers were quite familiar with them. The new tests use norms developed in 1988, so comparisons can be made to a population that is more reflective of today's students.

This is a particularly good time to change tests, because 1989-90 was the initial year of reform implementation. Last year, the Local School Councils were elected, organized, and trained; half of the principals were given performance contracts; and three-year improvement plans were adopted by all the schools. Introducing a new, more current test at this point permits us to establish a meaningful baseline against

which to measure progress of schools under the reform initiatives. Each year, as reform proceeds, we will be able to look at the results on the *ITBS* and the *TAP*, and compare them to the results in April 1990.

In changing tests and norms, however, it becomes difficult to compare the results this year to the performance of students and schools in previous years. While there are statistical methods of equating scores, the comparison becomes less valid when applied to the multiple changes in the norms and tests that have occurred since the Chicago Public Schools adopted the 1978 test.

As in the past, seventh grade ITBS scores were audited by the Department of Research, Evaluation and Planning. Students in 120 seventh-grade classrooms took a second test, using an alternative form, and given by staff from the department. The results were similar to previous audits: there was an average loss of three to four months. This may be due in part to motivation (students knew the tests did not affect their grades), in part because of differences in the test forms, and in part because of chance. The thirteen schools that showed substantial differences have been informed and will be included in next year's audit.

SUMMARY OF THE RESULTS ON THE ITBS

All elementary school students in the following groups were given the ITBS last spring:

- Students in the general program
- Students in bilingual education programs classified category C as of March 1, 1990 as well as limited-English-proficient category A and B students who were enrolled in a transitional bilingual education program for more than 3 years as of March 1, 1990
- Some students in special education -- behavior disordered, hard of hearing (only students who can be tested at Level 9 or above), moderately learning disabled, educable mentally handicapped, physically handicapped, other health impaired, and speech/language impaired

Testing is optional for other special education and limited-English proficient students.

There are two portions of the *ITBS* of most interest: Reading and Mathematics. Each of these tests contains subtests, examining different elements.

Reading Comprehension measures students' skills in three different areas. The first is facts or literal meaning: recognizing and understanding descriptions, classifications, sequence, and deducing the meanings of words or phrases from context. The second area is inference or interpretive meaning: interpreting the meaning of underlying relationships. This includes understanding cause and effect relationships, drawing conclusions, and inferring characters' traits, feelings, and motives. The third skill measured is generalization or evaluative meaning: developing general concepts from a specific selection of material.

Vocabulary measures both the size of the student's vocabulary and the ability to make fine distinctions in the meanings of words.

Mathematics skills are measured in three areas: the student's understanding of **Math Concepts**, that is, the quantitative process involved in mathematics; **Problem Solving**, which is the ability to work out problems using mathematical concepts;

and **Math Computation**, which is the ability to perform the various kinds of calculations required in mathematics. These subtests are summarized as the **Math Total** result, which is a composite score.

Summary Results

Tables 1 and 2, below summarize the results of the April 1990 *ITBS* testing in grades 1 through 8. Table 1 shows the number and percentage of students who scored at and above the national norms, and those who scored below the national norms, in each grade, for the most important of the tests: Reading Comprehension and Math Problem-Solving, along with the composite Math Total.

Table 1

STUDENTS ABOVE/BELOW NATIONAL NORMS, BY GRADE AND SUBTEST

THE *IOWA TESTS OF BASIC SKILLS* April 1990

Grade		Reading N	Reading %	Math Problem Solving N	Math Problem Solving %	Math Total N	Math Total %
1	At/above	6,389	25.6	9,589	38.4	9,386	37.5
	Below	18,609	74.4	15,409	61.6	15,612	62.52
2	At/above	6,957	26.8	8,377	32.2	9,564	36.8
	Below	19,040	73.2	17,620	67.8	16,433	63.23
3	At/above	5,397	19.6	8,317	30.2	7,802	28.3
	Below	22,158	80.4	19,238	69.8	19,753	71.74
4	At/above	5,442	17.9	7,982	26.2	7,792	25.6
	Below	25,011	82.1	22,471	73.8	22,661	74.45
5	At/above	6,723	23.4	7,560	26.3	7,186	25.0
	Below	22,064	76.6	21,227	73.7	21,601	75.06
6	At/above	6,027	22.8	7,755	29.3	7,558	28.6
	Below	20,436	77.2	18,708	70.7	18,905	71.47
7	At/above	7,760	30.3	6,945	27.1	6,963	27.2
	Below	17,868	69.7	18,683	72.9	18,655	72.88
8	At/above	7,299	29.7	7,157	29.1	7,046	28.7
	Below	17,261	70.3	17,403	70.9	17,514	71.3
Total	**At/above**	**51,994**	**24.2%**	**63,682**	**29.7%**	**63,297**	**29.5**
	Below	**162,447**	**75.8%**	**150,759**	**70.3%**	**151,144**	**70.5**

The information in this table is a useful baseline for measuring progress of the school system toward meeting the goal of the reform act: at least half of the students in each school will be at or above national norms.

In reading comprehension, about 1 in 4 students is doing as well as, or better than, the national norm. This ranges from a low of 17.9 percent in fourth grade to a high of 30.3 percent in seventh grade.

In solving math problems, almost 3 in 10 students reach or exceed the national norms. The low is 26.2 percent in fourth grade; the high is 38.4 percent in first grade. Overall math achievement is very close to these results: the average is 29.5 percent, with a low in fifth grade of 25 percent and a high in first grade of 37.5 percent.

These numbers suggest that the Chicago Public Schools will have to increase the proportion of children at or above norms by 20 to 25 percentage points to meet the goal of the reform plan.

Table 2

MEDIAN GRADE EQUIVALENT SCORES AND DIFFERENCES FROM NATIONAL NORMS ON THE *IOWA TESTS OF BASIC SKILLS*, April 1990

	READING						MATHEMATICS					
	Comprehension		Vocabulary		Concepts		Problem Solving		Computation		Total	
Gr.	Median	Dif.	Median	Dif.	Median	Dif.	Median	Dif.	Median	Dif.	Median	Dif.
1	1.4	-0.4	1.2	-0.6	1.6	-0.2	1.5	-0.3	1.8	0.0	1.6	-0.2
2	2.1	-0.7	2.1	-0.7	2.5	-0.3	2.3	-0.5	2.9	0.1	2.6	-0.2
3	2.9	-0.9	3.1	-0.7	3.3	-0.5	3.1	-0.7	3.3	-0.5	3.2	-0.6
4	3.8	-1.0	3.8	-1.0	4.0	-0.8	4.0	-0.8	4.4	-0.4	4.1	-0.7
5	4.8	-1.0	4.9	-0.9	4.9	-0.9	4.9	-0.9	5.3	-0.5	5.0	-0.8
6	5.8	-1.0	5.7	-1.1	5.9	-0.9	6.0	-0.8	6.2	-0.6	6.0	-0.8
7	7.1	-0.7	6.7	-1.1	6.8	-1.0	6.8	-1.0	7.5	-0.3	6.9	-0.9
8	8.1	-0.7	7.4	-1.4	7.9	-0.9	7.6	-1.2	8.1	-0.7	7.8	-1.0

Table 2 presents its information in *grade equivalent* scores.[1] This gives the median (middle ranking) score for all students taking the test, by grade. In addition, the

[1] A grade-equivalent number is expressed as a decimal, with the number to the left representing a year and the number to the right of the decimal indicating a month. Thus, a figure of 3.6 means grade 3, month 6 or the sixth month of the third grade. A grade equivalent score of, say, 6.5 means that the student performed as well on the test as we would expect a student at that grade and month to do. It does not mean that a student in a lower grade receiving this score is capable of doing sixth grade work.

differences from the norms are shown, also expressed as grade equivalents. These differences have been calculated on the basis of the grade plus .8, representing the fact that the test was given during the eighth month of instruction, e.g., 3.8 would be the expected median score for third grade.

Some observations can be made from the information in Table 2.

- Only in math computation do any median scores approach the national norms, and then only in first and second grades.
- In Reading Comprehension, the largest deficits are in the middle grades (4, 5, 6); in the upper grades, there is actually a reduction of the deficits.
- The gap between national norms and median grade-equivalent scores increases from the lower to the higher grades, suggesting that learning deficits are cumulative.

Although grade equivalents can be useful for tracking development over time, comparisons between grades and subjects are more readily performed when scores are expressed as percentile ranks. A percentile score indicates the percent of students nationally (based upon a sample) who scored below that figure. The national norm is 50, meaning that half of the students in the national sample did worse than, and half did better than, a student at the 50th percentile.

Table 3

MEDIAN PERCENTILE RANKS ON THE *IOWA TESTS OF BASIC SKILLS*

April 1990

	READING			MATHEMATICS		
Gr.	Comprehension	Vocabulary	Concepts	Problem Solving	Computation	Total
1	31	26	35	37	47	35
2	26	24	33	31	50	38
3	24	26	30	28	27	26
4	26	24	24	27	32	27
5	28	28	24	28	31	27
6	30	27	26	32	30	28
7	38	30	27	31	41	30
8	39	26	32	29	35	32

Several observations can be made from Table 3:

- Reading percentile ranks decline from first grade to third grade, but they increase steadily from grade 3 to grade 8.
- Math percentile ranks are highest in the first two grades; they decline in grade 3 and generally increase thereafter.
- The only median scores that reach or approach national norms are math computation for grades 1 and 2.

BREAKDOWNS OF *ITBS* RESULTS

The reform act emphasizes that the achievement of national norms is to occur despite students' race, ethnicity, gender and income status. To establish a baseline for this measure, the information in Table 1 has been broken down by gender and racial/ethnic identification.[2]

Table 4[3]
STUDENTS ABOVE/BELOW NATIONAL NORMS ON THE *ITBS*
BY GENDER AND RACIAL/ETHNIC BACKGROUND

		Reading		Problem Solving		Math Total	
		N	%	N	%	N	%
BOYS							
Asian	At/above	1,358	48.2	1,721	61.1	1,867	66.3
	Below	1,457	51.8	1,094	38.9	948	33.7
Black	At/above	11,386	17.5	15,608	24.0	14,939	23.0
	Below	53,630	82.5	49,408	76.0	50,077	77.0
Hispanic	At/above	5,186	21.0	7,340	29.7	7,043	28.5
	Below	19,496	79.0	17,342	70.3	17,639	71.5
Native Am.	At/above	83	40.5	96	46.8	88	42.9
	Below	122	59.5	109	53.2	117	57.1
White	At/above	5,799	44.4	6,998	53.6	6,779	52.0
	Below	7,248	55.6	6,049	46.4	6,268	48.0

[2] It is not yet possible to analyze individual student scores on the basis of income. Procedures are being developed to put this capability in place.

[3] Since the tests of some students could not be identified by gender, the totals in this table do not equal the totals in Table 1.

Table 4 (continued)

		Reading		Problem Solving		Math Total	
		N	%	N	%	N	%
GIRLS							
Asian	At/above	1,362	52.3	1,538	59.1	1,709	65.6
	Below	1,242	47.7	1,066	40.9	895	34.4
Black	At/above	15,100	22.3	17,150	25.4	17,475	25.8
	Below	52,508	77.7	50,458	74.6	50,133	74.2
Hispanic	At/above	5293	21.5	6,471	26.3	6,665	27.1
	Below	19304	78.5	18,126	73.7	17,932	72.9
Native Am.	At/above	91	38.1	88	36.8	95	39.7
	Below	148	61.9	151	63.2	144	60.3
White	At/above	6,260	47.4	6,562	49.7	6,533	49.5
	Below	6,936	52.6	6,634	50.3	6,663	50.5

As was pointed out previously, this information is intended to identify areas where added help is needed. It is not intended to imply that the achievement levels shown are caused by gender or racial/ethnic background. There is extensive research to show that neither gender or racial/ethnic background are accurate predictors of achievement. A better predictor is poverty, to the extent that it is indicative of social conditions that are impediments to learning.[4]

The figures in this table help to pinpoint areas where we have problems; generally black and Hispanic students -- both boys and girls -- are having greater difficulty in reading and math than are other students. As this type of information is provided for individual schools, it will help principals and teachers devise programmatic and instructional strategies targeted to the particular groups of students who are most in need of support.

[4] Poverty in itself is not an impediment to learning. Rather, it may indicate a social environment that hinders children in their pursuit of learning. In these cases, the schools must find alternative strategies to meet the academic needs of their students.

TESTS OF ACHIEVEMENT AND PROFICIENCY

The *Tests of Achievement and Proficiency (TAP)* is a nationally standardized norm-referenced achievement test, for use with high school students. It is a continuation of the *Iowa Tests of Basic Skills* that are given to elementary school students as part of the citywide testing program and it replaces the TAP, Form T, which was used from 1982 through 1988.

The TAP consists of six 40-minute subtests: Reading Comprehension, Mathematics, Written Expression, Using Sources of Information, Science, and Social Studies.

The **Reading Comprehension** subtest measures students' competence in reading texts in literature, social studies, and the sciences, and also materials encountered outside of school, such as advertisements, labels, and newspapers.

The primary purpose of the **Mathematics** subtest is to assess how well the student understands basic mathematical principles and how well he or she can apply these basic skills in everyday living. The areas covered are computation, knowledge of mathematical concepts, and problem solving.

The **Written Expression** subtest requires students to complete exercises in the mechanics of writing (spelling, capitalization, punctuation, grammar, and usage) as well as exercises in organizing, interrelating ideas, and selecting words for their proper placement within sentences.

The **Using Sources of Information** subtest is designed to measure students' skills in obtaining and interpreting information from a variety of sources, including maps, graphs, charts, tables, dictionaries, and books. Items deal with determining amounts and relationships, using data, and making inferences and generalizations.

The **Social Studies** subtest covers the areas of history, political science, geography, and other social sciences. Many questions involve the application of principles and generalizations rather than knowledge of specific facts.

The **Science** subtest includes items on biology, earth and space science, physics, and chemistry to measure students' understanding of scientific concepts, reasoning, and techniques.

In addition to the six subtests, students are also asked to complete a short questionnaire that deals with such issues as the degree to which they perceive their courses as valuable and enjoyable.

All high school students in the following groups were given the TAP:

- Students in a regular program (i.e., not receiving special education or bilingual services).
- Students in bilingual category B who have been enrolled in the system for at least three years, and students in bilingual category C. Testing is optional for others.
- Special education students who are behavior disordered, hard of hearing, moderately learning disabled, physically handicapped, other health impaired, and speech/language impaired. Testing is optional for others.

Table 5 summarizes the TAP results. It shows the number and percentage of students, by grade, who scored at or above the national norms in reading and math, as well as those who were below the national norms.

Table 5
STUDENTS ABOVE/BELOW NATIONAL NORMS ON THE *TAP*
By Grade and Subtest

Grade		Reading		Mathematics	
		N	%	N	%
9	at/above	7,772	29.6	5,127	19.5
	below	18,521	70.4	21,166	80.5
10	at/above	4,893	26.7	3,647	19.9
	below	13,458	73.3	14,704	80.1
11	at/above	4,686	27.7	3,365	19.9
	below	12,239	72.3	13,560	80.1
12	at/above	3,076	21.1	2,431	16.7
	below	11,514	78.9	12,159	83.3
Total	**at/above**	**20,427**	**26.8**	**14,570**	**19.1**
	below	**55,732**	**73.2**	**61,589**	**80.9**

The information in this table serves as a useful baseline for determining how much improvement must occur in each grade and subject to reach the legally required goal. The high schools are more successful in reading than in math, where less than 20 percent of the students scored at or above the national norms. In ninth grade reading, almost 3 of every 10 students scored at or above the norm; in twelfth grade math less than 2 of every 10 students did so.

The TAP results can also be reported in percentiles. A student's percentile score is the percent of students in a national sample who earned lower scores on the test. For example, a student with a percentile rank of 53 performed better than 53 percent of the students in the national norming sample. The national average is the fiftieth percentile.

Table 6 summarizes the results of the TAP by presenting the median percentile score for each grade, by subject.

Table 6
MEDIAN PERCENTILE RANKS ON THE *TAP*
(National norm is 50)

Grade	Reading Reading	Written Expression	Using Sources of Information	Mathematics	Science	Social Studies
9	36	38	33	20	39	33
10	31	30	33	24	37	40
11	29	38	34	22	33	36
12	33	33	31	24	26	28

It is difficult to make comparisons between this year's scores and those of previous years, both because of the changes in the tests and in the norms used.

BREAKDOWN OF *TAP* RESULTS

Table 7 presents the same information as in Table 5, broken down by gender and by racial/ethnic background. This does not imply, and is not intended to imply, that the achievement levels shown are caused by gender or by racial/ethnic background. The information is useful for diagnostic purposes: to identify particular groups of students who need additional assistance. No causal relationship is implied by the particular groupings used. In fact, these data point up the need for further study to determine the particular variables which do affect results -- so we can modify those factors and thereby improve student achievement.

Table 7[5]
STUDENTS ABOVE/BELOW NATIONAL NORMS ON THE *TAP*
By Gender and Racial/Ethnic Background

		READING		MATHEMATICS	
		N	%	N	%
BOYS					
Asian	at/above	725	43.3	1042	62.2
	below	951	56.7	634	37.8
Black	at/above	3706	17.9	2538	12.2
	below	17045	82.1	18213	87.8
Hispanic	at/above	2267	26.4	1775	20.7
	below	6315	73.6	6807	79.3
Native Am.	at/above	37	50.0	40	54.1
	below	37	50.0	34	45.9
White	at/above	2229	48.9	1934	42.4
	below	2330	51.1	2625	57.6
GIRLS					
Asian	at/above	689	43.7	893	56.6
	below	889	56.3	685	43.4
Black	at/above	5764	23.5	3142	12.8
	below	18788	76.5	21410	87.2
Hispanic	at/above	2475	27.6	1464	16.3
	below	6493	72.4	7504	83.7
Native Am.	at/above	32	51.6	23	37.1
	below	30	48.4	39	62.9
White	at/above	2420	51.9	1655	35.5
	below	2245	48.1	3010	64.5

[5]Since the tests of some students could not be identified by gender, the totals in this table do not equal those of Table 5.

INTERPRETING SCHOOL-LEVEL TEST RESULTS

Although standardized test information can be used for many purposes, its most important value is in helping local school councils, principals, and teachers to improve the program of instruction in the school. To do this, principals need to review and interpret the data on their students' performance.

Information Provided

To help schools in this process, the Department of Research, Evaluation and Planning provides schools with the following information:

- Summary of the **ITBS** results, by grade and subtest
- Summary of students above and below norm, by grade, by gender and by racial/ethnic background
- **Skills Analysis** for the **ITBS**, for each grade and subtest
- A **School Report** for the **IGAP**, prepared by the Illinois State Board of Education

This information can be used by principals, working with their Local School Councils and their teachers, to plan improvements in their educational programs.

Review with the Local School Council

The first step in the process is for the principal to discuss the results of the tests with the Local School Council. The purpose is to identify the areas of greatest need, so that the available resources can be targeted to address the most critical problems. This discussion should focus on two issues:

- How much progress has been made toward reaching national norms? How much progress toward the school improvement plan goals?
- In what areas has improvement occurred; in what areas has there been a decline?

Comparison to National Norms. The principal should review with the council the tables that show the percent of students scoring at or above the national norms. To get the best picture, the results for the last two or three years should be presented. (Because of the change in tests used, it is important to take the changes in norms into consideration.) That way, council members will be able to see where the improvements are occurring and how much progress is being made.

In looking at data from two or three years, it is useful to examine how a given grade did from year to year: for example, third grade in 1989 and in 1990. At the same time, it can be helpful to look at how the third grade class of 1989 did in 1990 as fourth graders. (Keep in mind, of course, that the membership of this class may have changed from one year to the next.)

In looking at the national norm comparisons, it is also a good idea to consider the information on gender and racial/ethnic background. Significant differences among the groups (e.g., boys vs. girls; blacks vs. whites) could highlight problem areas. In such cases, the council may need to closely study the instructional program to see how it can be made more effective with the lower achieving groups.

Grade Equivalent Summary. This chart provides the school with information that can help to pinpoint areas of concern. The following sample chart shows the kind of information presented:

Unit 1234
Sample School
District 12

CITYWIDE TEST RESULTS
IOWA TESTS OF BASIC SKILLS
GRADE-EQUIVALENT MEDIANS BY GRADE

GRADE	1	2	3	4	5	6	7	8
Students Tested	65	67	85	81	76	64	78	90
Vocabulary	1.7	2.6	3.3	3.9	5.1	5.7	7.1	8.4
Reading comprehension	1.6	2.5	3.2	3.8	5.2	5.9	7.5	8.4
Math concepts	1.6	2.6	3.3	3.9	5.1	6.5	6.8	7.9
Math problem solving	1.7	2.7	3.3	4.0	5.5	6.2	6.8	8.0
Math computation	2.0	2.9	3.3	4.3	5.1	6.3	7.4	8.4
Math total	1.7	2.5	3.4	4.1	5.0	6.4	7.0	8.1
Science					5.9		8.4	
Social studies					5.5		8.3	
Language skills					4.8		7.2	

The national norm grade-equivalent median score for each grade is the year in school plus eighth months, expressed thus: 5.8, meaning the fifth year (or grade) and

the eighth month (April, when the test is taken). By looking at the score shown on the summary and comparing it to the national norm, the principal can determine how far below, on average, his students in that grade performed last year. For example, in fifth grade reading comprehension, the students at this school had a median score of 5.2; they were .6, or 6 months, below the national norm.

Analysis of Results

In the example shown, the following conclusions could be made by the principal, in discussion with the council:

- Students in first and second grade are performing close to the national averages.
- Scores decline in third grade, where the average deficit is 5 months.
- Fourth grade has the largest average deficits: 8 months for each subtest.
- The largest deficits are in sixth grade vocabulary (11 months), fourth grade reading comprehension (10 months), and seventh grade math concepts and problem-solving (10 months each).
- Math concepts and problem-solving are seriously deficient in both seventh and eighth grades.
- Reading comprehension and vocabulary are particulary weak in the intermediate grades, 4 through 6.

Comparing these data with the results from the previous year, the principal and council can further focus on the areas of greatest need.

Using the information and analysis conducted, the principal and the Local School Council can decide, where to focus their efforts in improvement. For example, in the case shown above, the council and principal might decide to focus first on improving third grade, on the assumption that reducing or eliminating the deficits at this point would have the greatest long-term benefit. A secondary priority could be on math in the upper grades.

Which areas are to be targeted is a policy decision for the school community, involving the Local School Council, the principal, the Professional Personnel Advisory Committee, other advisory bodies, and the parents.

Identification of Problems

The process so far has served only to focus attention on the areas to be given priority. The next step is for the principal and teachers to identify the specific nature of the problems. This can be done through a careful analysis of the **Skills Analysis** sheets for the involved classrooms.

Specifying the Problem. For purposes of illustration, let us consider the problem of declining scores in third grade in the example above. The **Skills Analysis** sheet for the third grade in this school might look like this:

District: 12
UNIT: 1234
SCHOOL: SAMPLE

IOWA TESTS OF BASIC SKILLS

SKILLS ANALYSIS

Level: 09
Grade: 3
Room: 105

SKILLS	(C) No. of Items	(D) Average Number Tried	(E) Average Number Right	(F) % Right	(G) National % Right
READING COMPREHENSION	44	43	20	45	60
FACTS	20	20	10	50	60
Description/Categorization	14	14	8	57	63
Relationship/Context	6	6	2	33	55
INFERENCES	13	12	6	46	62
Cause & Effect/Conclusions	10	10	4	40	61
Traits & Feelings/Motives	3	3	1	33	67
GENERALIZATIONS	11	11	4	36	55
Main Idea/Organization	4	4	2	50	55
Application/Purpose	5	5	1	20	52
Language/Mood/Style	2	2	1	50	64

Looking at this table, the most revealing figures are in the two right-hand columns. These compare the average percentage of right answers in this third-grade class with the national average. It is apparent that the students were well below the national averages. They were best in 'FACTS' and weakest in 'GENERALIZATIONS'. They

also did poorly in 'INFERENCES.' Both 'Traits and Feelings/Motives' and 'Application/Purpose' are areas that need significant improvement.

Identifying Causes. The next step is to identify the specific causes or shortcomings in instruction that underly the problems. The principal may need to do one or more of the following:

- Review teacher's lesson plans to see how the problem areas are to be addressed. Also, to determine if the problem areas are taught before the test is given.
- Observe teacher's class to see how the problem areas are being taught.
- Analyze the instruction materials being used, to determine if the problem areas are being adequately covered.
- Review student attendance records and classroom management to determine if students are receiving the instruction being provided.

If the school has more than one third grade class, the review of both **Skills Analysis** sheets might indicate that one classroom performed significantly better (or worse) than the other. In this case, the principal would need to look at how students are grouped. If the rooms are equally mixed, the principal would then need to compare the classes, to determine why one is more successful than the other.

Finding Solutions

The final step for the principal is to determine the solution to the problem. In the example being shown, the principal would have several choices:

- If the lesson plans show the subjects are not being covered, or are inappropriately being covered, the principal can work with the teacher to revise the lesson plans.
- If students are not attending, special activities can be instituted to improve attendance; parents should be involved in this effort.

- If the instruction materials are inadequate, they should be replaced or supplemented.

- If the teacher is not effectively presenting the material (e.g., the questioning of students relies too heavily on yes/no answers), some staff development is called for. The extent and formality of this would depend upon the extensiveness of the need. It could consist of a brief discussion, possibly with other teachers, or could necessitate a formal program.

Whatever intervention is chosen, it should be implemented as soon as possible. Next year, the principal of this school could look at the 1991 scores to see if the 1990-91 third grade class performed better than the 1989-90 class did. This will indicate the effectiveness of the intervention taken this year.

Summary

The process described here would need to be repeated, to address all the priority areas of the school. Also, all teachers should be expected to review the **Skills Analysis** sheets for their classes and should seek ways to improve the teaching of the specific skills identified as weakest.

In analyzing the test results, principals and councils should be careful about drawing comparisons between schools and -- within the school -- between classrooms. There are many factors that can affect test results; improper comparisons can be both misleading and damaging. The emphasis should be placed instead upon an analysis of deficits from the norms and the steps to be taken to overcome these deficits.

FUTURE REPORTING OF TEST RESULTS

The information in this report, as mentioned before, serves as a baseline for assessing the progress of students in the Chicago Public Schools under the impetus of the reform initiatives. Reports in the coming years will look back to this report, to identify increases in the number and percentage of students who have reached or exceeded national norms on the ITBS and the TAP.

Additional information will be analyzed in the future reports, focusing on gains made by students from year to year. Examining gains provides an early measure of the progress being made by schools. This can be done by analyzing grade-equivalent scores and percentile ranks on a comparative basis, from year to year. For example, by subtracting this year's grade four grade-equivalent score in reading comprehension from next year's grade five score, it will be possible to measure how much progress students made over the year. Similarly, with percentile ranks, it is possible to compare the year-to-year scores to determine if schools are making progress.

The gains analysis can be done for the school system as a whole and for individual schools. At the systemwide level, average gains will be reported by grade. At the school level, average gains will be reported by groups of grades: primary (1 - 3), intermediate (4 - 6) and upper (7-8). This degree of aggregation is necessary to reduce the distortions that could occur in results if the number of students being averaged is too small.

Beyond the analysis of gains, the proportion of students scoring in each quarter will be looked at as an indication of changes occurring within the top and bottom halves of the test results.[6] This information is currently reported on the state's report cards for all schools in Illinois.

As an example of this type of data, the following information is included in the report cards issued in October, for ITBS results in grades 3, 6 and 8:

---

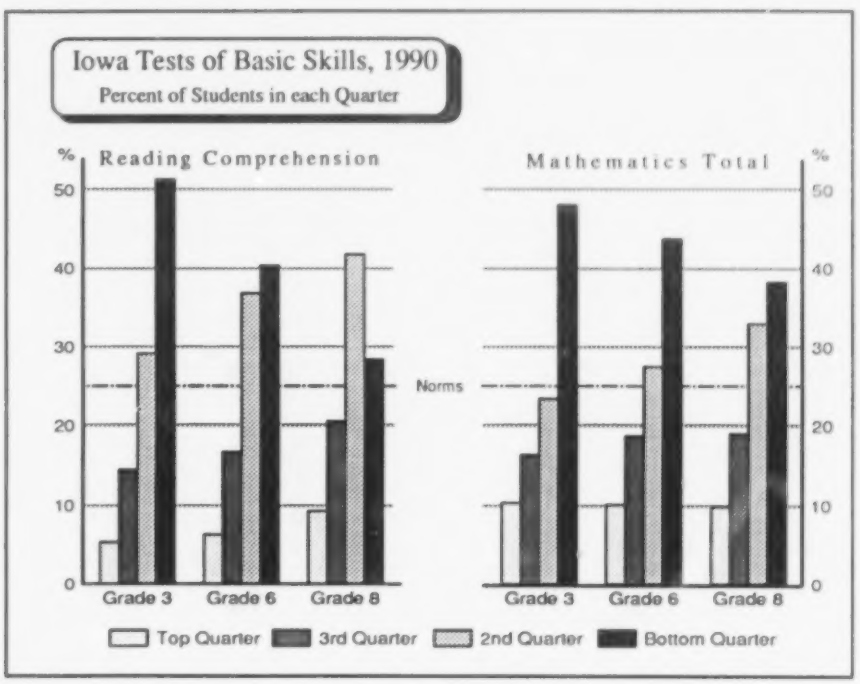
[6] In a normal distribution of results, the scores of 25 percent of students will fall into each quarter. Negative variations from that distribution (too many students in the bottom quarters, too few in the top quarters) indicate the extent of the disparity that schools must address.

Table 8

PERCENT OF STUDENTS IN EACH QUARTER ON
THE IOWA TESTS OF BASIC SKILLS, 1990

	Reading Comprehension			Mathematics Total		
	Gr. 3	Gr. 6	Gr. 8	Gr. 3	Gr. 6	Gr. 8
Top Quarter	5.3	6.3	9.3	10.3	10.1	9.9
Third Quarter	14.3	16.7	20.5	18.3	18.7	19.0
Second Quarter	29.1	36.8	41.8	23.4	27.5	32.9
Bottom Quarter	51.3	40.3	28.4	48.0	43.7	38.2

This same information can be presented graphically, which shows the deviation from the normal distribution.



Adding the top two quarters indicates the percentage of students above norms; adding the bottom two indicates the percentage below. Thus, in grade 3 reading comprehension, 19.6 percent of the students were at or above the norm, while 80.4 percent were in the bottom half.

The significance of the quarters is that the extent of the problem is more readily understood. In this case, for instance, over half of all students in grade 3 scored in the bottom quarter on reading comprehension. So, in order to increase the percent of students above the norm from 19.6 percent to over 50 percent, the schools will have to address the needs of students who are more than a little below norms. By contrast, in grade 8 reading comprehension 29.8 percent (9.3 + 20.5) of the students are already above the norm. In order to achieve the goal of 50 percent, an additional 20.1 percent need to move up. Since the second quarter contains 41.8 percent of the students, the goal should be attainable without intensive intervention.

The use of quarters also highlights the educational needs of students who are significantly below the norm. A key objective of school improvement is to raise the achievement levels of all students. Thus, positive changes in the quarters will indicate the degree to which all students' learning needs are being successfully addressed.

In future reports, data will be presented for all grades. The analysis of progress will focus both on the percentage of students above the norms, and on the distribution of students' scores within the quarters. Particular note will be made of schools with large proportions of students in the bottom quarter, where intensive intervention may be appropriate.